

Proposal for characters for Khowar, Torwali, and Burushaski

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: **Proposal to add characters needed for Khowar, Torwali, and Burushaski**

2. Requester's name: *(1) Elena Bashir, Ph.D.; (2) Sarmad Hussain, Ph.D.; (3) Deborah Anderson, Script Encoding Initiative, UC Berkeley*

3. Requester type (Member body/Liaison/Individual contribution): *(1) Individual researcher at University of Chicago; (2) Head, Center for Research in Urdu Language Processing, FAST National University of Computer and Emerging Science, Lahore, with proposal supported by National Bodies and local user-community organizations; (3) Liaison member SC2 and WG2*

4. Submission date: *May 2006*

5. Requester's reference (if applicable): *Dept. of South Asian Languages and Civilizations, University of Chicago (USA)*

6. Choose one of the following:
 This is a complete proposal: *Complete proposal*
 (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
 a. This proposal is for a new script (set of characters):
 Proposed name of script: _____
 b. The proposal is for addition of character(s) to an existing block: *Yes*
 Name of the existing block: *Arabic Block (U0600) or Arabic Supplement Block (U075-077F)*

2. Number of characters in proposal: *16 precomposed characters*

3. Proposed category (select one from below - see section 2.2 of P&P document):
 A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection)
 C-Major extinct D-Attested extinct E-Minor extinct
 F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols

4. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): *3*
 Is a rationale provided for the choice? *Yes*
 If Yes, reference: *These characters are part of the basic alphabets of the languages concerned (see appendices).*

5. Is a repertoire including character names provided? *Yes*
 a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*
 b. Are the character shapes attached in a legible form suitable for review? *Yes*

6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? *Dr. Sarmad Hussain, Center for Research in Urdu Language Processing, National University of Computer and Emerging Sciences, B Block Faisal Town, Lahore, Pakistan (sarmad.hussain@nu.edu.pk)*
 If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: _____

7. References:
 a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *yes*
 b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *yes (see appendices)*

8. Special encoding issues:
 Does the proposal address other aspects of character data processing (if applicable) such as input,

¹ Form number: N3002-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10)

presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

no

(except that usual roman representations are given for the sounds involved)

9. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	<i>no</i>
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom?	<i>Yes</i> <i>(1) Center for Research in Urdu Language Processing, National University of Science and Technology, Lahore, Pakistan. (2) National Language Authority, Islamabad, Pakistan. (3) Anjuman-e-Taraqqi, Khowar. (4) Burushaski Research Academy (5) Mr. Inam Ullah, Village Bahrain, Swat, NWFP., representing Torwali Educational, Literary and Linguistic Society.</i>
If YES, available relevant documents:	<i>See document L2/06-150.</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>Yes</i> <i>See sections 2,3, and 4 in the proposal.</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>Common</i> <i>See sections 2, 3, and 4.</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>yes</i> <i>Examples from published works included. (see appendices)</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	<i>Yes</i> <i>Yes</i> <i>In modern use</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>Not necessarily</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>No</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>Yes.</i> <i>Yes</i> <i>See section 2</i>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>No</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	<i>No</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	<i>No</i>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	<i>No</i>

Proposal

1. Introduction

This proposal introduces 16 characters necessary for writing the languages Khowar, Torwali, and Burushaski in such a way as to (i) utilize the script and characters approved and in use by the language using communities, and (ii) represent the sounds of these languages accurately. Khowar and Torwali are NW Indo-Aryan languages spoken in the North West Frontier Province of Pakistan, while Burushaski (linguistic isolate) is spoken in the Northern Areas of Pakistan.

The considerations and issues involved are discussed separately for each of these three languages.

2. Khowar. There are approximately 300,000 speakers of Khowar, most of whom live in the Chitral District, where Khowar is the lingua franca of the local area. Khowar is written in a modified Perso-Arabic script, using the Nastaliq style. These characters were devised to represent the retroflex sibilant and affricate sounds of Khowar first in 1921 by Prince Nasir-ul-Mulk in his booklet *Khowar Qaida*, and reinforced in a later *Khowar Qaida* by his nephew Prince Samsam-ul-Mulk (sample page included, Appendix 2). These characters have been in consistent use since that time. There is an active literary community in Chitral, and to date approximately 40 books have been published in Khowar. A pilot dictionary for Khowar is underway and is located at: <http://dsal.uchicago.edu/dictionaries/khowar>. The "front matter" section of this digital dictionary discusses in detail the issue of the lack of Unicode characters. Samples of published documents including these characters are included in this proposal, Appendices 1 and 2.

Four characters necessary for Khowar are not included in the Unicode Standard (Arabic block). These four are in Table 1.

Since the UCS Arabic model involves precomposition, the proposed characters below (as well as those for Torwali and Burushaski, below) are all pre-composed.

Table 1: Proposed characters required for Khowar

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Conventional Roman rep.
1	ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH BELOW	voiced retroflex affricate	Made up of ARABIC LETTER HAH (062D) and ARABIC LETTER TAH (0637). For suggested contextual forms, see Table 3. In initial and medial forms, the small letter TAH is below the letter HAH. The Khowar name is Jiim. (See Appendices 1 and 2 for examples [circled])	j (with underdot)
2	ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH AND TWO DOTS	voiceless retroflex affricate	Made up of ARABIC LETTER HAH (062D), ARABIC LETTER TAH (0637), and two dots. The Khowar name is Ce. (See Appendices 1 and 2 for examples [in squares])	ç

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Conventional Roman rep.
3	ARABIC LETTER SEEN WITH SMALL ARABIC LETTER TAH AND TWO DOTS	voiceless retroflex sibilant	Made up of ARABIC LETTER SEEN (0633), ARABIC LETTER TAH (0637) and two dots. The Khowar name is Se (Şe). The name of this character appears sometimes as Siin, (Şiin); however most speakers refer to it as Se (Şe). (See Appendices 1 and 2 for examples [in rectangles])	Ş
4	ARABIC LETTER REH WITH SMALL ARABIC LETTER TAH AND TWO DOTS	voiced retroflex sibilant	Made up of ARABIC LETTER REH (0631), ARABIC LETTER TAH (0637) and two dots. The Khowar name is Ze. (See Appendices 1 and 2 for examples [in ovals])	ž

3. Torwali. Torwali is spoken by about 60,000 people in and around Village Bahrein in Swat, NWFP (Pakistan). A pilot dictionary for Torwali is underway and is located at: <http://dsal.uchicago.edu/dictionaries/torwali/>. The "front matter" section of this digital dictionary discusses in detail the issue of the lack of Unicode characters. As of the time of writing this proposal, one character necessary for writing Torwali remains to be added to the Unicode standard and is shown in Table 2 below. Please see Table 3 for examples of the contextual forms of this character and the Khowar character.









Appendix 3 shows the character in Inam Ullah's section on "Torwali language and literature," from the M.Phil. student's textbook, *Pashto, Hindko, Torwali, Gawri* (2004). Appendix 4 is a screenshot of the Torwali dictionary page with a special font for the retroflex character.

Table 2: Proposed character required for Torwali

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Conventional Roman rep.
5	ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH ABOVE	voiced retroflex affricate	Made up of ARABIC LETTER HAH (062D) and ARABIC LETTER TAH (0637). For suggested contextual forms, see Table 3. (See Appendices 3 and 4.)	j (with underdot)

Proposed contextual forms for this character (in contrast with the Khowar ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH BELOW) are provided in the table below.

Table 3: Suggested contextual forms for Torwali and Khowar characters (#1 and #5) for retroflex j

Letter	Isolated form	Initial form	Medial form	Final form
Torwali				
Khowar				



* Note that the initial and medial Khowar forms have the small letter *below* the base form, whence the name.

4. Burushaski. According to one recent estimate, there are approximately 200,000 speakers of Burushaski, living mostly in the Hunza Valley in the Northern Areas of Pakistan, but with sizeable numbers of the community also in Gilgit, Rawalpindi/Islamabad, and Karachi.





Literary publication in Burushaski has been going on for about 20 years, using the Perso-Arabic script (essentially Urdu script) modified with additional characters necessary for writing Burushaski sounds and phonological features not present in Urdu. This Burushaski alphabet was devised by Allama Nasir ud Din Nasir Hunzai, the universally acknowledged dean of Burushaski letters. It first appeared in the booklet *Basic Burushaski*, published in 1984 by the Burushaski Research Committee in Hunza, Gilgit and Karachi. The recent project of compiling a multi-volume Burushaski dictionary, which is currently underway in Karachi, is being carried out by the Burushaski Research Academy using the alphabet established by Allama Nasir ud Din Hunzai. Sample pages from vol. 1 of the recent dictionary and other works are included with this proposal (see Appendix 5).

The characters needed for Burushaski are listed and described in Table 4. The proposed names for the Burushaski characters are those recommended by the Burushaski Research Academy.

Table 4. Proposed characters required for Burushaski

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Suggested vernacular name
 6	ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.	short [a], usually stressed	Made up of ARABIC LETTER ALEF (0627) and EASTERN ARABIC-INDIC DIGIT TWO (06F2). (See Appendix 5-1 for example)	alif-i-maqsurah
 7	ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE.	extra long [a], (with low-rising tone)	Made up of ARABIC LETTER ALEF (0627) and EASTERN ARABIC-INDIC DIGIT THREE (06F3). (See Appendix 5-2 for example)	alif-a-zul hijjatayn

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Suggested vernacular name
8	ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.	[i] in final position (usually stressed).	Made up of ARABIC LETTER FARSI YEH (06CC) and EASTERN ARABIC-INDIC DIGIT TWO (06F2). This character does not appear in word-initial position. In letter-group (ligature)-initial position, it appears with two dots below and the numeral two above (identical to #13). In medial position it appears with two dots below and Arabic diacritic two above (identical to #13). In final position it has diacritic two above, but no dots below. Thus it differs from #13 only in final position. Table 5 shows all the forms based on YEH. (See Appendix 5-3 for examples)	ya-i maqsurah
9	ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE.	long [i] (with high tone)	Made up of ARABIC LETTER FARSI YEH (06CC) and EASTERN ARABIC-INDIC DIGIT THREE (06F3). Table 5 shows all the contextual forms based on YEH. (See Appendix 5-4 for example)	ya-i zul hijjatayn
10	ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW.	velar approximant	Made up of ARABIC LETTER FARSI YEH (06CC) and EASTERN ARABIC-INDIC DIGIT FOUR (06F4). Represented by \dot{y} in literature on Burushaski. (See Appendix 6 for example)	yeh
11	ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.	[o] short o, often stressed.	Made up of ARABIC LETTER WAW (0648) and EASTERN ARABIC-INDIC DIGIT TWO (06F2). (See Appendix 5-5 for example)	waw-i majhulah-i maqsurah
12	ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE.	long [o]	Made up of ARABIC LETTER WAW (0648) and EASTERN ARABIC-INDIC DIGIT THREE (06F3). (See Appendix 5-6 for example)	waw-i zul hijjatayn

Glyph Ref. No.	Description and suggested Unicode name	Phonetic characterization	Comments	Suggested vernacular name
 13	ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.	short [e], (often stressed)	Made up of ARABIC LETTER YEH BARREE (06D2) and EASTERN ARABIC-INDIC DIGIT TWO (06F2). This character does not appear in word-initial position. In letter-group (ligature)-initial position, it appears with two dots below and the numeral two above (identical to #8). In medial position, it is identical to ARABIC YEH with superscript 2 (#8). Thus it differs from #8 only in final position. Table 5 shows all the contextual forms based on YEH. (See Appendix 5-7 for example)	ya-i majhulah-i-maqsurah
 14	ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE	long e (with high tone?)	Made up of ARABIC LETTER YEH BARREE (06D2) and EASTERN ARABIC-INDIC DIGIT THREE (06F3). Table 5 shows all the contextual forms based on YEH. (See Appendix 5-8 for example)	yae-i mujhulah-i-maqsurah
 15	ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW	voiceless aspirated dental affricate	Made up of ARABIC LETTER HAH (062D) and EASTERN ARABIC-INDIC DIGIT FOUR (06F4). Conventional Roman representation is [çh]. Table 6 shows all the contextual forms. In initial and medial forms, the digit is below the letter HAH. (See Appendices 5-9, 6a*, 6b and 7 for examples.)	che
 16	ARABIC LETTER SEEN WITH EASTERN ARABIC-INDIC DIGIT FOUR ABOVE	voiceless retroflex fricative	Made up of ARABIC LETTER SEEN (0633) and EASTERN ARABIC-INDIC DIGIT FOUR (06F4). Conventional Roman representation is [ʂ]. (See Appendices 5-10, 6a*, 6b and 7 for examples)	se

* Note: The third column in Appendix 6a (from *Asqúriine Basí*) points out the similarity of the Burushaski sounds with sounds in Shina, a language that uses four dots to represent retroflexion. Even in the earliest publication of the alphabet, a work entitled *Basic Burushaski*, four dots are not used (see Appendix 7). The examples from Appendix 5 (5-9 and 5-10) provide further evidence that only characters with numerals appear in *Burushaski lughat* [dictionary], the most recent and authoritative publication. Hence, there is no reason to view character 16 as a variant of U+075C ARABIC LETTER SEEN WITH FOUR DOTS ABOVE,

Table 5: Isolated, Initial, Medial and Final forms of characters based on YEH

Letter	Isolated form	Initial form	Medial form	Final form
#8 ya-i maqsurah	ی	ی	ی	ی
#13 ya-i majhulah-i- maqsurah	ے	یر	یر	ے
#9 ya-i zul hijjatayn	ی	یر	یر	ی
#14 yae-i mujhulah-i maqsurah	ے	یر	یر	ے
ARABIC LETTER FARSI YEH (Unicode 06CC)	ی	یر	یر	ی
ARABIC LETTER YEH BARREE (Unicode 06D2)	ے	یر	یر	ے

Note: Proposed letters #8 and #13 behave in the same way as their base Arabic characters (Unicode 06CC and Unicode 06D2). That is, they differ only in their isolated and final forms. Note that #9 and #14 also follow this pattern, which is already established for Unicode 06CC and 06D2 (best known for their use in Urdu).

Table 6: Isolated, Initial, Medial and Final forms of ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW (character #15)

Letter	Isolated form	Initial form	Medial form	Final form
ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW	ح	ح	ح	ح

Note that the initial and medial forms have the small letter *below* the base form, whence the name.

4. Character Properties

All characters have the general category of "Lo."

5. Summary. The additional characters thus proposed are:

For Khowar:

U+076E ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH BELOW
 U+076F ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH AND TWO DOTS
 U+0770 ARABIC LETTER SEEN WITH SMALL ARABIC LETTER TAH AND TWO DOTS
 U+0771 ARABIC LETTER REH WITH SMALL ARABIC LETTER TAH AND TWO DOTS

For Torwali:

U+0772 ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH ABOVE

For Burushaski:

U+0773 ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE
 U+0774 ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE
 U+0775 ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE
 U+0776 ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE
 U+0777 ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW
 U+0778 ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE
 U+0779 ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE
 U+077A ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE
 U+077B ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE
 U+077C ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW
 U+077D ARABIC LETTER SEEN WITH EASTERN ARABIC-INDIC DIGIT FOUR ABOVE

APPENDICES

These appendices include supporting materials for the proposal in the form of scanned photocopies of examples of published works using the proposed characters. These documents are listed here, and then included below.

Khowar Examples

Appendix 1: Rahmat ud Din and Daud, 1989. *Phuphukan Kitaab*. Chitral: Anjuman e-Taraqqi Khowar
 Appendix 2: Samsam ul Mulk. n.d. *Khowar Qaida*. Peshawar: Idaaraa e Ishaa'at, Sarhad.

Torwali Examples

Appendix 3: Chapter by Inam Ullah on "Torwali language and literature," from *Pashto, Hindko, Torwali, Gawri*, ed. Inamul Haq Javed, 2004. [Torwali section of a textbook with language materials on Pashto, Hindko, Torwali, and Gawri for M.Phil. students]
 Appendix 4: Screenshot of Digital Dictionary of Torwali page with special font used for retroflex character (<http://dsal.uchicago.edu/cgi-bin/romadict.pl?query=a&display=utf8&table=torwali>)

Burushaski Examples

Appendix 5: Burushaski Research Academy. In press. *Burushaski lughat* [Burushaski dictionary]. Karachi: Burushaski Research Academy.
 Appendix 6a and 6b: Hunzai, Nasiruddin Nasir. n.d. *Asqúriñe Basí* [Garden of blossoms/roses]. Karachi: Burushaski Research Academy.
 Appendix 7: Burushaski Research Committee. 1985. *Basic Burushaski*. Hunza, Gilgit, Karachi: Burushaski Research Committee (=Burushaski Research Academy)

Bibliographies**Bibliography for Khowar**

Rahmat ud Din and Daud, 1989. *Phuphukan Kitaab*. Chitral: Anjuman e-Taraqqi Khowar
 Samsam ul Mulk. n.d. *Khowar Qaida*. Peshawar: Idaaraa e Ishaa'at, Sarhad.

Bibliography for Torwali

Karimi, Abdul Hameed. *Kohistani Urdu Bolchal*.

Torwali, Zuber. *Torwali Alif Be*.

_____. *Ish-o-maash* (A Bear and Man). [a short story]

_____. *Laash* (Honor). [a short story]

Ullah, Inam. "Torwali language and literature." In *Pashto, Hindko, Torwali, Gawri*, ed. Inamul Haq Javed. Islamabad: Department of Pakistani Languages, Allama Iqbal Open University Islamabad, 2004. [Torwali section of language materials for M.Phil students]

_____. *Torwali-English-Urdu dictionary*. (in preparation)

_____. *Torwali-Urdu lughaat*. (in preparation)

Bibliography for Burushaski

Burushaski Research Academy. In press. *Burushaski lughat* [Burushaski dictionary]. Karachi: Burushaski Research Academy.

Burushaski Research Committee. 1985. *Basic Burushaski*. Hunza, Gilgit, Karachi: Burushaski Research Committee (=Burushaski Research Academy).

Hunzai, Nasiruddin Nasir. n.d. *Asqúriñe Basí*. [Garden of blossoms/roses] Karachi: Burushaski Research Academy.

Shaheen, S.M. et al. (eds.) *Burushaski jawaahar paare* ["Jewels of quicksilver", selected writings of Nasiruddin Nasir Hunzai]. Karachi: Burushaski Research Academy, Karachi University, and Pakistan Academy of Letters, Karachi.

Letters of Support

Document L2/06-150 includes four letters of support from concerned language-user groups.

کھوارو الف، ب

س	سین	(اگ)	آنکار
ش	شوم	(بجائی)	برابر
ص	شپیک	(بکری)	پائے
ض	صابون	(آبا)	تت
ط	منورت	(بجرا)	ٹیپ
ظ	طولی	(ثواب)	ثوابی
ع	ظالم	(داماد)	جار
غ	عدت	(زنجیر)	چینچیر
ف	فرخ	(مٹکی)	چوٹ
ق	قف	(اشک)	چوچھو
ک	کومورو	(چوٹے)	ضیق
گ	کان	(حرام)	حرام
ل	لو	(پنچکی)	خرابا
م	میک	(اکڑی)	دار
ن	نان	(لڈکا)	ڈق
و	واؤ	(ذہن)	ذہن
ہ	ہوست	(گتا)	رینی
ع	بی	(لوڑی)	ڑوڑ
ی	یور	(پھاڑ)	زوم
ے	غیٹھ	(بیٹا)	واؤ
		(پتلا)	ڑوغ

(جسارت پنڈت زبیر گلروڈ۔ لاہور)

حرف ادبے ہیٹان نام

الف	ا	ب	پ	ت	ط	ث
بیم	بج	بچ	بچ	بچ	بچ	بچ
خ	د	ڈ	ذ	ر	ڑ	ڑ
ز	ژ	ڑ	س	س	س	س
ص	ص	ط	ظ	ع	ع	ع
ف	ق	ک	گ	ل	ل	ل
ن	و	ہ	ہ	ی	ی	ی

ہدایت

چونکہ کھوار زبان پہلے بار سکولوں میں رائج ہو رہی ہے، اسلئے اساتذہ ماجد کے عرض ہے کہ وہ طلباء کو پڑھانے سے پہلے غور و تبادلی کیا کریں کھوار کے مخصوص حروف اچھی طرح پہچان لیں اور طلباء کو بھی ذہن نشین کریں حروف حسب ذیل ہیں:-

حرف	نام حرف	کھوار لفظ	اردو ترجمہ
(۱) طح	طحیم	طخیر	رنجیر
(۲) شخ	شخے	شخ	چھوٹا بچہ
(۳) طح	طحے	چھوٹا بچہ	نمبر دار
(۴) طز	تزے	زانگ	اوپا
(۵) شش	ششین	شا	کالا

حرف "ٹ" کھوار میں بھی موجود ہے مگر اس حرف کی آواز اردو کے "ڈ" سے تھوڑی مختلف ہے کھوار میں اس حرف کو "ول" اور "ڈ" کے بین میں ادا کیا جاتا ہے۔ ان حروف کے علاوہ اس کا خیال رکھا جائے کہ کھوار قواعد کے اصولوں کے مطابق اسما کے ساتھ مختلف حروف لگ کر ان کی اصل شکل کو تبدیل کر دیتے ہیں۔ یہ اضافت یا حالت مفعولی یا حرف جار وغیرہ کے مخصوص حروف ہیں۔

خالص کھوار الفاظ میں حروف ط - ح - ذ - ض - ض - ط - ظ - ع استعمال نہیں ہوتے صرف ان الفاظ میں جو کسی دوسری زبان سے لئے گئے ہوں یہ حروف استعمال کئے جاسکتے ہیں:-

از مؤلف

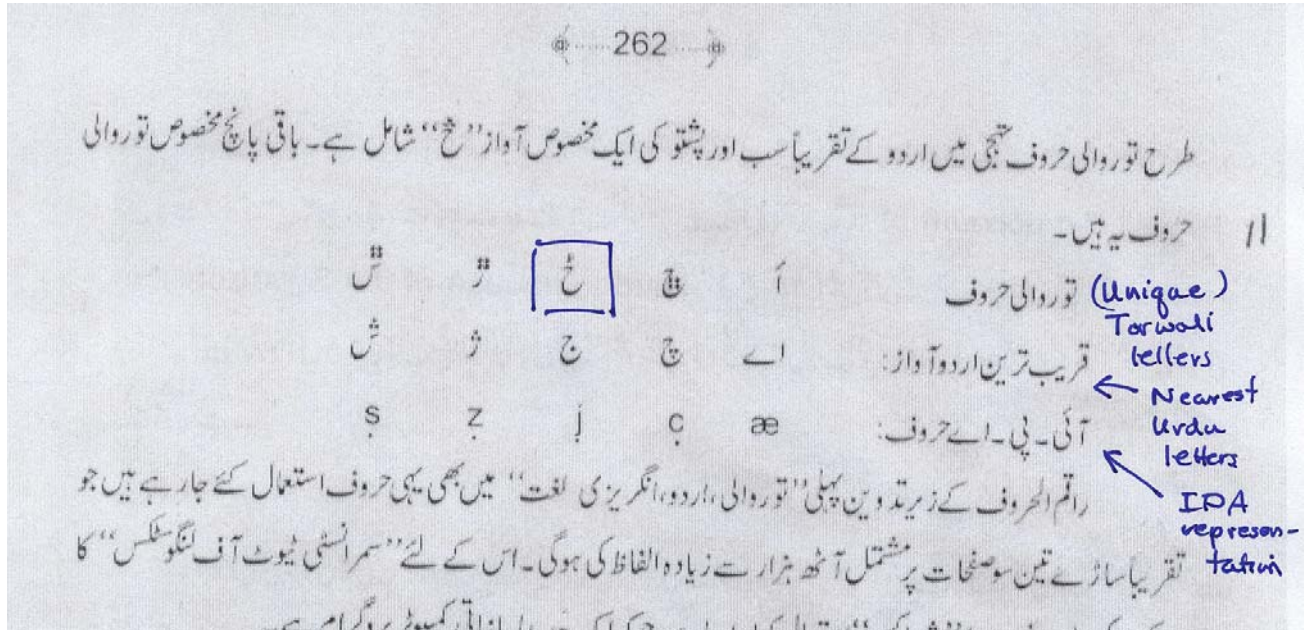
بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

اب پ ت ٹ ش ج چ
 ح ڄ ڇ ڙ ڙ ڙ ڙ ڙ
 س ش ش ص ض ط ط ط
 غ ف ق ک گ ل م ن

و ۵ ۶ ی کے

پچھتراری زبان عام استعمال حروف صرف پچھتراری
 فنی کے موزار پچھتراری زبانوں خاص پچھتراری استاؤ
 وقان تے نشان دیار

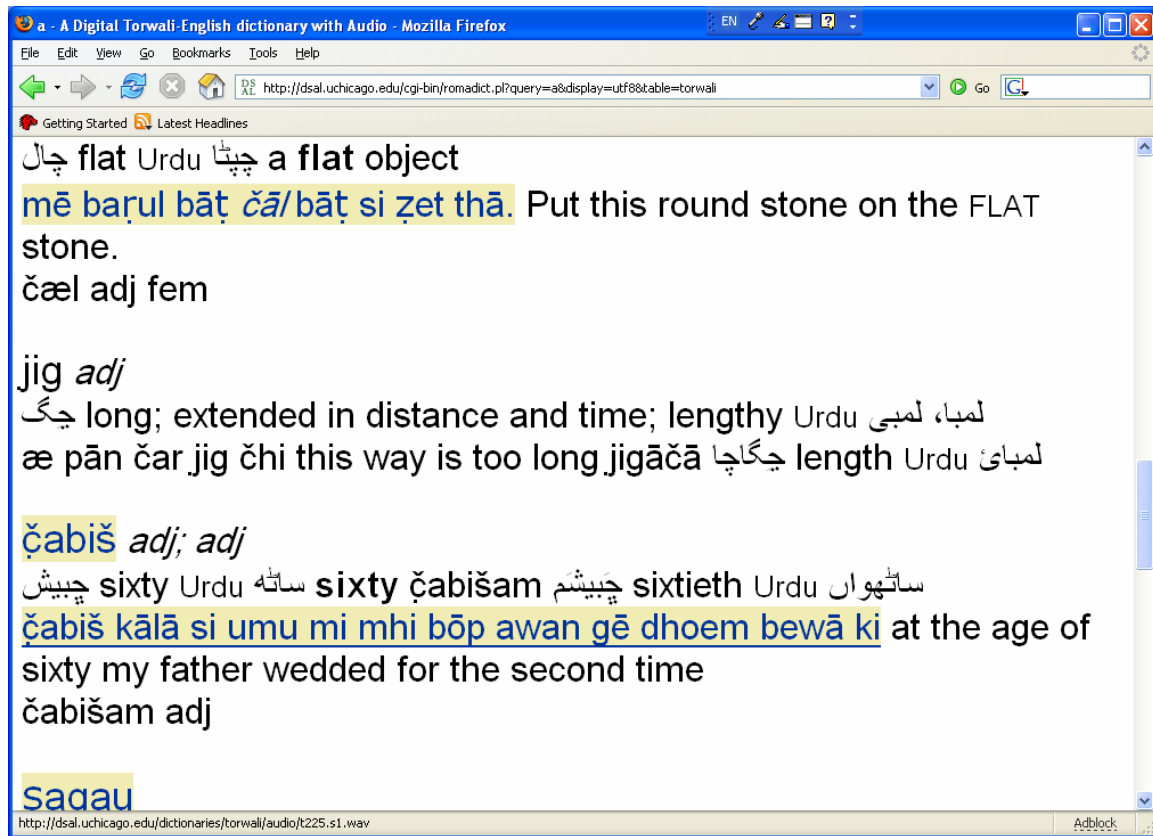
Appendix 3: Sample of Torwali letter ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH ABOVE (character #5)



Text from chapter on "Torwali language and literature" by Inam Ullah, showing Torwali letter character #5 in print. (From the book *Pashto, Hindko, Torwali, Gawri* ed. Inamul Haq Javed, Islamabad, 2004).

Appendix 4: Torwali ARABIC LETTER HAH WITH SMALL ARABIC LETTER TAH ABOVE (character #5)

Screenshot of Digital Dictionary of Torwali page illustrating the problem faced in the absence of this character and the temporary solution adopted for the voiced retroflex affricate, represented by j[dot below].



Screenshot of original entry for j[dot below]ig 'long' from Inam Ullah's dictionary database, in which SIL-made character for this letter is used.

Appendix 5: Examples of Burushaski letters from *Burushaski lughat* [Burushaski dictionary].

5-1: Evidence for character #6, ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE

ب (ا)

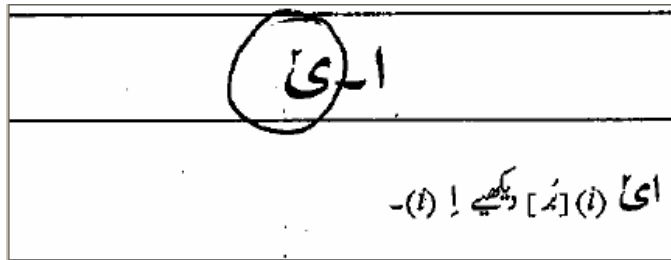
با (ba) [بُر۔ اَم] بوسہ۔
 با ایتس (ba étas) [بُر] بوسدینا، چومنا۔
 بائپس (bappas) [بُر۔ اَم] محل، قصر۔
 دینے بائپسے پگھلنے تک منے و آنصیر
 تاکہ پیتشما دینے آقا سیرٹیک
 ترجمہ: اے نصیر! قصر دین کے دروازے کی خاک بن جانا کرٹو۔
 دین کے مالک کے اسرار کا مشاہدہ کر سکے گا۔

5-2 : Evidence for character #7, ARABIC LETTER ALEF WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE

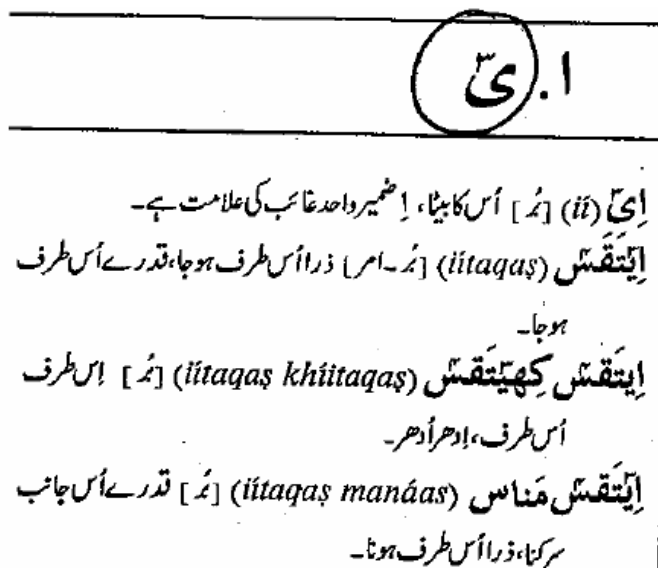
آ

آ (aa) یہ الف ممدودہ کی ایک شکل ہے جو صرف بُرُوشسکی لہجہ کے ساتھ مخصوص ہے، الف مدہ اور اس کی آواز کا فرق صرف بُرُوشسکی لہجے میں معلوم ہوتا ہے۔
 آپو (aápo) [بُر] کم تھوڑا، قلیل۔
 آپو آپو کے تھڑا پو (aápo aápo ke thadaápo)
 [بُر۔ کہاوت] لالچ بری بلا ہے۔
 آپور تھڑا پو (aápuwar thadaápo) [بُر۔ کہاوت] لالچ

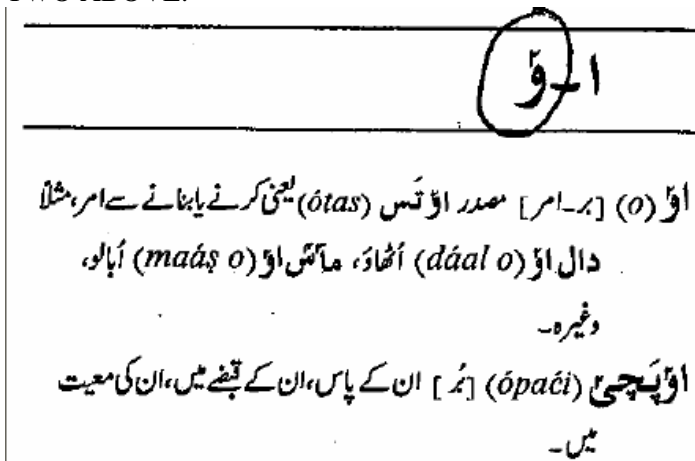
5-3: Evidence for character #8, ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE



5-4: Evidence for character #9, ARABIC LETTER FARSI YEH WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE



5-5: Evidence for character #11, ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.



5-6: Evidence for character #12, ARABIC LETTER WAW WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE.

(۱۳)

اَوّ (oó) [بر] میرے آنسو (۲) یہ لفظ بعض کلمات کے شروع میں علامت نفی کے طور پر بھی آتا ہے اور نئی کے معنی بھی دیتا ہے۔

اَوّت (oóʔ) [بر] نقصان، خسارہ، زیاں۔

اَوّدی (oódi) [بر-اسم] اس لفظ کا پہلا جز ”اَوّ“ حرف نفی ہے اور دوسرا جز ”دی“ ”دوہنا تھن سے دودھ نکالنا“ پس اَوّدی (oódi) کے معنی ہیں کہ تم اپنی بکریوں کو مت دوہنا۔ تفصیل کے لیے برڈ شسکی کتاب برڈ شسکی جواہر پارے ”ہونزا کے روایتی تہوار قسط۔۲“

5-7: Evidence for character #13, ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT TWO ABOVE.

(۱۲)

ایے (e) [بر-امر] کر لے، بنا لے، یہ لفظ بعض افعال کے ساتھ امر میں بطور لاحقہ بھی آتا ہے، مثلاً آر ایے (ar e) = ڈرالے، پہت ایے (phat e) = چھوڑ دے، زپ ایے (zap e) = یاد کر۔

وغیرہ۔

5-8 : Evidence for character #14, ARABIC LETTER YEH BARREE WITH EASTERN ARABIC-INDIC DIGIT THREE ABOVE.

(۱۳)

ایے (eé) [بر] حرف نفی والے نکار، یہ بعض مصادر مفعول اور افعال امر پر داخل ہو کر نہیں بنتی جاتا ہے مثلاً دیکس (délas) سے ایڈیلکس (éedilas) اور ایڈیل (éedil) وغیرہ۔

5-9: Evidence for character #15, ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW



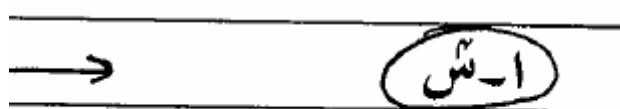
اِحْأ (tchâ) [ہ] دیکھیے اِحْأیس (ichâyas) جس کا پیار

-ہے

اِحْأپن (ichâpan) [ہ] دیکھیے اِحْأپنس (ichâpanas)

جس کا پیار ہے۔

5-10 Evidence for character #16, ARABIC LETTER SEEN WITH EASTERN ARABIC-INDIC DIGIT FOUR ABOVE



اَس (as) [ہ-اسم] میری گردن، ا ضمیر واحد شکلم کی علامت ہے۔

اَس (is) [ہ-اسم] تھکی، پیار کی مار۔

اَس اَس ایتس (is is éetas) [ہ-مجاورہ] روتے ہوئے بچوں کو

جدید حروف کی صوتی مثالیں

ترجمہ	لفظ	علامت	جدید حروف	شمار
سیدھا	چن	∴ = ۴	چ	۱
بھائی	اخوا	∴	خ	۲
سنتری	ڈر	∴	ڈ	۳
بیزار	ژاو	∴	ژ	۴
شاخ	سگر	∴ = ۴	گس	۵
ہمیشہ	ضا	∴	ض	۶
نیند	دگ	∴	گ	۷
میرے باپ	ایا	∴ = ۴	یا	۸

Appendix 6a: Burushaski examples from Nasiruddin Nasir Hunzai. The table shows examples of the proposed characters (from top to bottom) #15, #16, and #10.

Note: The column with four dots equated to the numeral for "4" are given in order to point out the similarity of the Burushaski sounds with sounds in Shina, a language that uses four dots to represent retroflexion. See Appendices 6b and 7 for other examples with the numeral "4."

اُردو سے زائد

ژ	خ	چ
ض	س	ث
-	ی	گ

نوٹ: جن حروف کے اندر یا اوپر یا نیچے ۴ (چار) کی علامت لگی ہے وہ چار نقطوں کے معنی میں ہے، یعنی بجائے اس کے کہ ہم اوپر کے تین حروف کو اس طرح لکھیں: چ، ث، ی، ان کو یوں آسانی سے لکھ دیتے ہیں: چ، س، ی۔

نصیر الدین نصیر (حب علی) ہونزائی

Appendix 6b: Examples of Burushaski characters #15 and #16 from a listing of the alphabet by Nasiruddin Nasir Hunzai (*Asqúriine Bast*). Note the use of the numeral.

بُروشسکی حرپُڈ

ترجمہ	لفظ	حرپ	ترجمہ	لفظ	حرپ
ھما	خلک	خ	اُونٹ	اُٹ	ا
سینا	شک	خ	بیلی	بُش	ب
مالک	دمن	د	صُراحی	پُون	پ
ڈھول	ڈڈنگ	ڈ	ہڈی	تِن	ت
ذات، خاندان	ذات	ذ	سالم	ٹوآک	ٹ
کالی چڑیا	ڈیو/ ڈو	ڈ	ثواب	ثباپ	ث
تیار	رل	ر	سلام	جُو	ج
سڑک	سڑک	ڑ	چپانا	چق	چ
ہلنا/ ہلانا	زل	ز	جلدی	چوڑ، چوڑ	چ
نمبردار	ڑشگگین	ڑ	حرام	حرام	ح

ترجمہ	لفظ	حرپ	ترجمہ	لفظ	حرپ
نیند	دنگ	گ	کرامات	سرمگ	س
شادی	گر	گ	سج	شوآقم	ش
گلینہ	لیس	ل	سبزنگ	سقم رگ	س

Appendix 7: Examples of Burushaski characters #15 and #16 from *Basic Burushaski* by the Burushaski Research Committee (1985). Note the use of the numeral.

Samples of Burushaski character #15 (ARABIC LETTER HAH WITH EASTERN ARABIC-INDIC DIGIT FOUR BELOW)

پندول	٤ پيرو	٤
گيند	٣ طھري	٣
توت	توت	٤
جھجان	جھ	٣
چسراخ	چسرق	٣
مچھلی	٣ مچھو	٣
پانی	٣ پل	٣
		٣

Sample of Burushaski character #16 (ARABIC LETTER SEEN WITH EASTERN ARABIC-INDIC DIGIT FOUR ABOVE)

٣

سڑک	سڑک	٣
زود	زود	٣
گدھا	٣ گھن	٣
نمرگوش	سر	٣
غنچہ	٣ گھنکلا	٣
روٹی	٣ ٹیک	٣
		٣